Chapter 5

Linear Systems

Few physical elements display truly linear characteristics. For example the relation between force on a spring and displacement of the spring is always nonlinear to some degree. The relation between current through a resistor and voltage drop across it also deviates from a straight-line relation. However, if in each case the relation is ?reasonably? linear, then it will be found that the system behavior will be very close to that obtained by assuming an ideal, linear physical element, and the analytical simplification is so enormous that we make linear assumptions wherever we can possibly to so in good conscience.

R. Cannon, *Dynamics of Physical Systems*, 1967 [Can03].

In Chapters 2–4 we considered the construction and analysis of differential equation models for physical systems. We placed very few restrictions on these systems other than basic requirements of smoothness and well-posedness. In this chapter we specialize our results to the case of linear, time-invariant, input/output systems. This important class of systems is one for which a wealth of analysis and synthesis tools are available, and hence it has found great utility in a wide variety of applications.

5.1 Basic Definitions

We have seen several examples of linear differential equations in the examples of the previous chapters. These include the spring mass system (damped oscillator) and the operational amplifier in the presence of small (non-saturating) input signals. More generally, many physical systems can be modeled very accurately by linear differential equations. Electrical circuits are one example of a broad class of systems for which linear models can be used effectively. Linear models are also broadly applicable in mechanical engineering, for example as models of small deviations from equilibria in solid and fluid mechanics. Signal processing systems, including digital filters of the sort used in CD and MP3 players, are another source of good examples, although often these are best modeled in discrete time (as described in more detail in the exercises).

In many cases, we *create* systems with linear input/output response through the use of feedback. Indeed, it was the desire for linear behavior that led Harold S. Black, who invited the negative feedback amplifier, to the principle of feedback as a mechanism for generating amplification. Almost all modern single processing systems, whether analog or digital, use feedback to produce linear or near-linear input/output characteristics. For these systems, it is often useful to represent the input/output characteristics as linear, ignoring the internal details required to get that linear response.

For other systems, nonlinearities cannot be ignored if one cares about the global behavior of the system. The predator prey problem is one example of this; to capture the oscillatory behavior of the coupled populations we must include the nonlinear coupling terms. However, if we care about what happens near an equilibrium point, it often suffices to approximate the nonlinear dynamics by their local *linearization*, as we already explored briefly in Section 4.3. The linearization is essentially an approximation of the nonlinear dynamics around the desired operating point.

Linearity

We now proceed to define linearity of input/output systems more formally. Consider a state space system of the form

$$\begin{aligned} \frac{dx}{dt} &= f(x, u) \\ y &= h(x, u), \end{aligned} \tag{5.1}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^p$ and $y \in \mathbb{R}^q$. As in the previous chapters, we will usually restrict ourselves to the single input, single output case by taking $p = q = 1$. We also assume that all functions are smooth and that for a reasonable class of inputs (e.g., piecewise continuous functions of time) that the solutions of equation (5.1) exist for all time.

It will be convenient to assume that the origin $x = 0$, $u = 0$ is an equilibrium point for this system ($\dot{x} = 0$) and that $h(0, 0) = 0$. Indeed, we can do so without loss of generality. To see this, suppose that $(x_e, u_e) \neq (0, 0)$ is an equilibrium point of the system with output $y_e = h(x_e, u_e) \neq 0$. Then we can define a new set of states, inputs, and outputs

$$\tilde{x} = x - x_e \qquad \tilde{u} = u - u_e \qquad \tilde{y} = y - y_e$$

and rewrite the equations of motion in terms of these variables:

$$\begin{aligned} \frac{d}{dt}\tilde{x} &= f(\tilde{x} + x_e, \tilde{u} + u_e) &&=: \tilde{f}(\tilde{x}, \tilde{u}) \\ \tilde{y} &= h(\tilde{x} + x_e, \tilde{u} + u_e) - y_e &&=: \tilde{h}(\tilde{x}, \tilde{u}). \end{aligned}$$

In the new set of variables, we have that the origin is an equilibrium point with output 0, and hence we can carry our analysis out in this set of variables. Once we have obtained our answers in this new set of variables, we simply have to remember to "translate" them back to the original coordinates (through a simple set of additions).

Returning to the original equations (5.1), now assuming without loss of generality that the origin is the equilibrium point of interest, we write the output $y(t)$ corresponding to initial condition $x(0) = x_0$ and input $u(t)$ as $y(t; x_0, u)$. Using this notation, a system is said to be a *linear input/output system* if the following conditions are satisfied:

(i) $y(t; \alpha x_1 + \beta x_2, 0) = \alpha y(t; x_1, 0) + \beta y(t; x_2, 0)$

(ii) $y(t; \alpha x_0, \delta u) = \alpha y(t; x_0, 0) + \delta y(t; 0, u)$

(iii) $y(t; 0, \delta u_1 + \gamma u_2) = \delta y(t; 0, u_1) + \gamma y(t; 0, u_2)$.

Thus, we define a system to be linear if the outputs are jointly linear in the initial condition response and the forced response. Property (ii) is the usual decomposition of a system response into the homogeneous response ($u = 0$) and the particular response ($x_0 = 0$). Property (iii) is the formal definition of the the *principle of superposition* illustrated in Figure 5.1.

Example 5.1 (Scalar system)**.** Consider the first order differential equation

$$\begin{aligned} \frac{dx}{dt} &= ax + u \\ y &= x \end{aligned}$$

with $x(0) = x_0$. Let $u_1 = A \sin \omega_1 t$ and $u_2 = B \cos \omega_2 t$. The homogeneous solution the ODE is

$$x_h(t) = e^{at} x_0$$

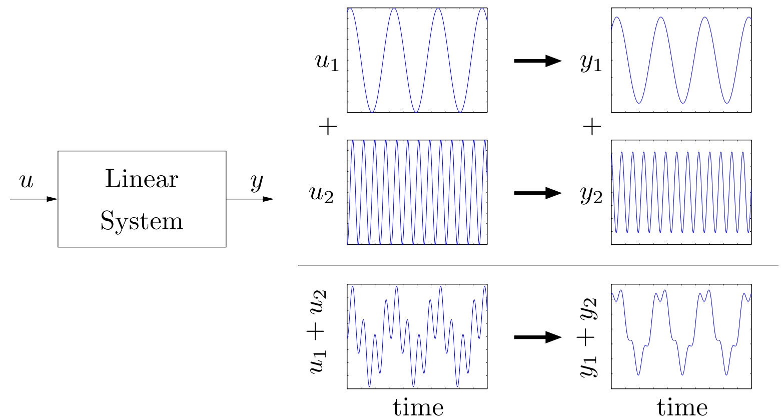


Figure 5.1: Illustration of the principle of superposition. The output corresponding to $u_1 + u_2$ is the sum of the outputs y_1 and y_2 due to the individual inputs.

and the two particular solutions are

$$x_1(t) = -A\frac{-\omega_1 e^{at} + \omega_1 \cos\omega_1 t + a \sin\omega_1 t}{a^2 + \omega_1^2}$$
$$x_2(t) = B\frac{a e^{at} - a\cos\omega_2 t + \omega_2 \sin\omega_2 t}{a^2 + \omega_2^2}.$$

Suppose that we now choose $x(0) = \alpha x_0$ and $u = u_1 + u_2$. Then the resulting solution is

$$x(t) = e^{at}\left(\alpha x(0) + \frac{A\omega_1}{a^2+\omega_1^2} + \frac{Ba}{a^2+\omega_2^2}\right) - A\frac{\omega_1\cos\omega_1 t + a\sin\omega_1 t}{a^2+\omega_1^2} + B\frac{-a\cos\omega_2 t + \omega_2\sin\omega_2 t}{a^2+\omega_2^2} \tag{5.2}$$

(to see this, substitute the equation in the differential equation). Thus, the properties of a linear system are satisfied for this particular set of initial conditions and inputs. ∇

We now consider a differential equation of the form

$$\frac{dx}{dt} = Ax + Bu$$
$$y = Cx + Du, \tag{5.3}$$

where $A \in \mathbb{R}^{n\times n}$ is a square matrix, $B \in \mathbb{R}^n$ is a column vector of length n, C is a row vector of width n and D is a scalar. (In the case of a multi-input systems, B, C and D becomes a matrices of appropriate dimension.)

Equation (5.3) is a system of linear, first order, differential equations with input u, state x and output y. We now show that this system is a linear input/output system, in the sense described above.

Proposition 5.1. *The differential equation* (5.3) *is a linear input/output system.*

Proof. Let $x_{h1}(t)$ and $x_{h2}(t)$ be the solutions of the linear differential equation (5.3) with input $u(t) = 0$ and initial conditions $x(0) = x_{01}$ and x_{02}, respectively, and let $x_{p1}(t)$ and $x_{p2}(t)$ be the solutions with initial condition $x(0) = 0$ and inputs $u_1(t), u_2(t) \in \mathbb{R}$. It can be verified by substitution that the solution of equation (5.3) with initial condition $x(0) = \alpha x_{01} + \beta x_{02}$ and input $u(t) = \delta u_1 + \gamma u_2$ and is given by

$$x(t) = \big(\alpha x_{h1}(t) + \beta x_{h2}(t)\big) + \big(\delta x_{p1}(t) + \gamma x_{p2}(t)\big).$$

The corresponding output is given by

$$y(t) = \big(\alpha y_{h1}(t) + \beta y_{h2}(t)\big) + \big(\delta y_{p1}(t) + \gamma y_{p2}(t)\big).$$

By appropriate choices of α, β, δ and γ, properties (i)–(iii) can be verified. □

As in the case of linear differential equations in a single variable, we define the solution $x_h(t)$ with zero input as the *homogeneous* solution and the solution $x_p(t)$ with zero initial condition as the *particular* solution. Figure 5.2 illustrates how these the homogeneous and particular solutions can be superposed to form the complete solution.

It is also possible to show that if a system is input/output linear in the sense we have described, that it can always be represented by a state space equation of the form (5.3) through appropriate choice of state variables.

Time Invariance

Time invariance is another important concept that is can be used to describe a system whose properties do not change with time. More precisely, if the input $u(t)$ gives output $y(t)$, then if we shift the time at which the input is applied by a constant amount a, $u(t + a)$ gives the output $y(t + a)$. Systems that are linear and time-invariant, often called LTI systems, have the interesting property that their response to an arbitrary input is completely characterized by their response to step inputs or their response to short "impulses".

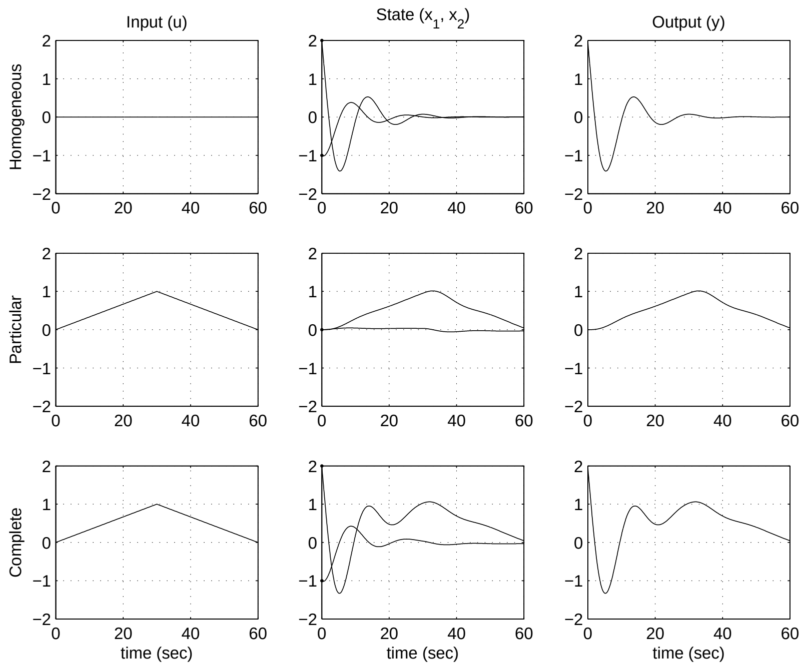


Figure 5.2: Superposition of homogeeous and particular solutions. The first row shows the input, state and output corresponding to the initial condition response. The second row shows the same variables corresponding to zero initial condition, but nonzero input. The third row is the complete solution, which is the sume of the two individual solutions.

We will first compute the response to a piecewise constant input. Assume that the sytem is initially at rest and consider the piecewise constant input shown in Figure 5.3a. The input has jumps at times t_k and its values after the jumps are $u(t_k)$. The input can be viewed as a combination of steps: the first step at time t_0 has amplitude $u(t_0)$, the second step at time t_1 has amplitude $u(t_1) - u(t_0)$, etc.

Assuming that the system is initially at an equilibrium point (so that the initial condition response is zero), the response to the input can then be obtained by superimposing the responses to a combination of step inputs. Let $H(t)$ be the response to a unit step applied at time t. The response to the first step is then $H(t - t_0)u(t_0)$, the response to the second step is

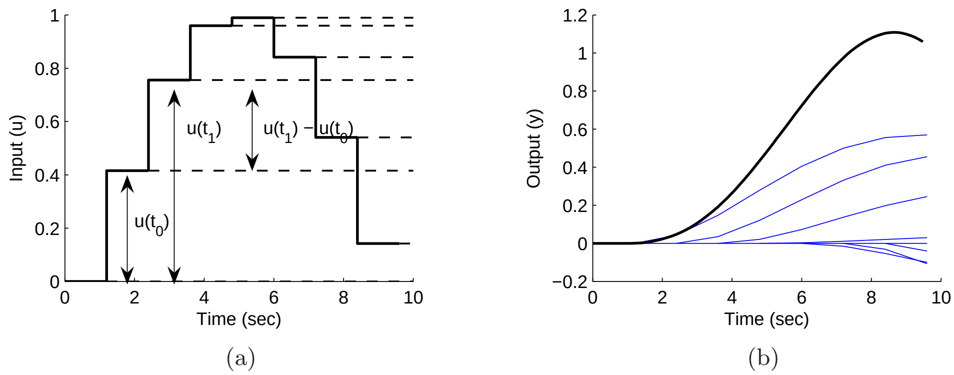


Figure 5.3: Response to piecewise constant inputs: (a) a piecewise constant signal can be represented as a sum of step signals; (b) the resulting output is the sum of the individual outputs.

$H(t-t_1)\big(u(t_1)-u(t_0)\big)$, and we find that the complete response is given by

$$\begin{aligned} y(t) &= H(t-t_0)u(t_0) + H(t-t_1)\big(u(t_1)-u(t_0)\big) + \cdots \\ &= \big(H(t)-H(t-t_1)\big)u(t_0) + \big(H(t-t_1)-H(t-t_2)\big)u(t_1) \\ &= \sum_{n=0}^{\infty}\big(H(t-t_n)-H(t-t_{n+1})\big)u(t_n) \\ &= \sum_{n=0}^{\infty}\frac{H(t-t_n)-H(t-t_{n+1})}{t_{n+1}-t_n}\big(t_{n+1}-t_n\big)u(t_n). \end{aligned}$$

An example of this computation is shown in Figure 5.3b.

The response to a continuous input signal is obtained by taking the limit as $t_{n+1}-t_n \to 0$, which gives

$$y(t) = \int_0^{\infty} H'(t-\tau)u(\tau)d\tau, \tag{5.4}$$

where H' is the derivative of the step response, which is also called the *impulse response*. The response of a linear time-invariant system to any input can thus be computed from the step response. We will derive equation (5.4) in a slightly different way in the next section.

5.2 The Convolution Equation

Equation (5.4) shows that the input response of a linear system can be written as an integral over the inputs $u(t)$. In this section we derive a more general version of this formula, which shows how to compute the output of a linear system based on its state space representation.

The Matrix Exponential

Although we have shown that the solution of a linear set of differential equations defines a linear input/output system, we have not fully computed the solution of the system. We begin by considering the homogeneous response corresponding to the system

$$\frac{dx}{dt} = Ax. \tag{5.5}$$

For the *scalar* differential equation

$$\dot{x} = ax \qquad x \in \mathbb{R}, a \in \mathbb{R}$$

the solution is given by the exponential

$$x(t) = e^{at}x(0).$$

We wish to generalize this to the vector case, where A becomes a matrix.

We define the *matrix exponential* as the infinite series

$$e^X = I + X + \frac{1}{2}X^2 + \frac{1}{3!}X^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!}X^k, \tag{5.6}$$

where $X \in \mathbb{R}^{n\times n}$ is a square matrix and I is the $n \times n$ identity matrix. We make use of the notation

$$X^0 = I \qquad X^2 = XX \qquad X^n = X^{n-1}X,$$

which defines what we mean by the "power" of a matrix. Equation (5.6) is easy to remember since it is just the Taylor series for the scalar exponential, applied to the matrix X. It can be shown that the series in equation (5.6) converges for any matrix $X \in \mathbb{R}^{n\times n}$ in the same way that the normal exponential is defined for any scalar $a \in \mathbb{R}$.

Replacing X in equation (5.6) by At where $t \in \mathbb{R}$ we find that

$$e^{At} = I + At + \frac{1}{2}A^2t^2 + \frac{1}{3!}A^3t^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!}A^kt^k,$$

and differentiating this expression with respect to t gives

$$\frac{d}{dt}e^{At} = A + At + \frac{1}{2}A^3t^2 + \cdots = A\sum_{k=0}^{\infty}\frac{1}{k!}A^kt^k = Ae^{At}. \tag{5.7}$$

Multiplying by $x(0)$ from the right we find that $x(t) = e^{At}x(0)$ is the solution to the differential equation (5.5) with initial condition $x(0)$. We summarize this important result as a theorem.

Theorem 5.2. *The solution to the homogeneous system of differential equation* (5.5) *is given by*

$$x(t) = e^{At}x(0).$$

Notice that the form of the solution is exactly the same as for scalar equations.

The form of the solution immediately allows us to see that the solution is linear in the initial condition. In particular, if x_{h1} is the solution to equation (5.5) with initial condition $x(0) = x_{01}$ and x_{h2} with initial condition x_{02}, then the solution with initial condition $x(0) = \alpha x_{01} + \beta x_{02}$ is given by

$$x(t) = e^{At}\big(\alpha x_{01} + \beta x_{02}\big) = \big(\alpha e^{At}x_{01} + \beta e^{At}x_{02}\big) = \alpha x_{h1}(t) + \beta x_{h2}(t).$$

Similarly, we see that the corresponding output is given by

$$y(t) = Cx(t) = \alpha y_{h1}(t) + \beta y_{h2}(t),$$

where y_{h1} and y_{h2} are the outputs corresponding to x_{h1} and x_{h2}.

We illustrate computation of the matrix exponential by three examples.

Example 5.2 (Double integrator)**.** A very simple linear system that is useful for understanding basic concepts is the second order system given by

$$\begin{aligned}\ddot{q} &= u\\ y &= q.\end{aligned}$$

This system system is called a *double integrator* because the input u is integrated twice to determine the output y.

In state space form, we write $x = (q, \dot{q})$ and

$$\frac{dx}{dt} = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} x + \begin{pmatrix}0\\ 1\end{pmatrix} u.$$

The dynamics matrix of a double integrator is

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

and we find by direct calculation that $A^2 = 0$ and hence

$$e^{At} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

Thus the homogeneous solution ($u = 0$) for the double integrator is given by

$$\begin{aligned} x(t) &= \begin{pmatrix} x_1(0) + tx_2(0) \\ x_2(0) \end{pmatrix} \\ y(t) &= x_1(0) + tx_2(0). \end{aligned}$$

∇

Example 5.3 (Undamped oscillator)**.** A simple model for an oscillator, such as the spring mass system with zero damping, is

$$m\ddot{q} + kq = u.$$

Putting the system into state space form, the dynamics matrix for this system is

$$A = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & 0 \end{pmatrix}$$

We have

$$e^{At} = \begin{pmatrix} \cos\omega_0 t & \frac{1}{\omega_0}\sin\omega_0 t \\ -\omega_0 \sin\omega_0 t & \cos\omega_0 t \end{pmatrix} \qquad \omega_0 = \sqrt{\frac{k}{m}},$$

and the solution is then given by

$$x(t) = e^{At}x(0) = \begin{pmatrix} \cos\omega_0 t & \frac{1}{\omega_0}\sin\omega_0 t \\ -\omega_0 \sin\omega_0 t & \cos\omega_0 t \end{pmatrix} \begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix}.$$

This solution can be verified by differentiation:

$$\begin{aligned} \frac{d}{dt}x(t) &= \begin{pmatrix} -\omega_0 \sin\omega_0 t & \cos\omega_0 t \\ -\omega_0^2 \cos\omega_0 t & -\omega_0 \sin\omega_0 t \end{pmatrix} \begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix}. \\ &= \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & 0 \end{pmatrix} \begin{pmatrix} \cos\omega_0 t & \frac{1}{\omega_0}\sin\omega_0 t \\ -\omega_0 \sin\omega_0 t & \cos\omega_0 t \end{pmatrix} \begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix} = Ax(t). \end{aligned}$$

If the damping c is nonzero, the solution is more complicated, but the matrix exponential can be shown to be

$$e^{At} = e^{\frac{-ct}{2m}} \begin{pmatrix} \dfrac{e^{\omega_d t} + e^{-\omega_d t}}{2} + \dfrac{e^{\omega_d t} - e^{-\omega_d t}}{2\sqrt{c^2 - 4km}} & \dfrac{e^{\omega_d t} - e^{-\omega_d t}}{\sqrt{c^2 - 4km}} \\ -\dfrac{ke^{\omega_d t} - ke^{-\omega_d t}}{\sqrt{c^2 - 4km}} & \dfrac{e^{\omega_d t} + e^{-\omega_d t}}{2} - \dfrac{ce^{\omega_d t} - ce^{-\omega_d t}}{2\sqrt{c^2 - 4km}}, \end{pmatrix}$$

where $\omega_d = \sqrt{c^2 - 4km}/2m$. Note that ω_d can either be real or complex, but in the case it is complex the combinations of terms will always yield a positive value for the entry in the matrix exponential. ∇

Example 5.4 (Diagonal system)**.** Consider a diagonal matrix

$$A = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

The kth power of At is also diagonal,

$$(At)^k = \begin{pmatrix} \lambda_1^k t^k & & & 0 \\ & \lambda_2^k t^k & & \\ & & \ddots & \\ 0 & & & \lambda_n^k t^k \end{pmatrix}$$

and it follows from the series expansion that the matrix exponential is given by

$$e^{At} = \begin{pmatrix} e^{\lambda_1 t} & & & 0 \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ 0 & & & e^{\lambda_n t} \end{pmatrix}.$$

∇

Eigenvalues and Modes

The initial condition response of a linear system can be written in terms of a matrix exponential involving the dynamics matrix A. The properties of the matrix A therefore determine the resulting behavior of the system. Given a matrix $A \in \mathbb{R}^{n\times n}$, recall that λ is an eigenvalue of A with eigenvector v if λ and v satisfy

$$Av = \lambda v.$$

In general λ and v may be complex valued, although if A is real-valued then for any eigenvalue λ, its complex conjugate λ^* will also be an eigenvalue (with v^* as the corresponding eigenvector).

Suppose first that λ and v are a real-valued eigenvalue/eigenvector pair for A. If we look at the solution of the differential equation for $x(0) = v$, it follows from the definition of the matrix exponential that

$$e^{At}v = \big(I + At + \frac{1}{2}A^2t^2 + \cdots\big)v = (v + \lambda t v + \frac{\lambda^2 t^2}{2}v + \cdots\big)v = e^{\lambda t}v.$$

The solution thus lies in the subspace spanned by the eigenvector. The eigenvalue λ describes how the solution varies in time and is often called a *mode* of the system. If we look at the individual elements of the vectors x and v, it follows that

$$\frac{x_i(t)}{x_j(t)} = \frac{v_i}{v_k},$$

and hence the ratios of the components of the state x are constants. The eigenvector thus gives the "shape" of the solution and is also called a *mode shape* of the system.

Figure 5.4 illustrates the modes for a second order system. Notice that the state variables have the same sign for the slow mode $\lambda = -0.08$ and different signs for the fast mode $\lambda = -0.62$.

The situation is a little more complicated when the eigenvalues of A are complex. Since A has real elements, the eigenvalues and the eigenvectors are complex conjugates

$$\lambda = \sigma \pm j\omega \quad \text{and} \quad v = u \pm jw,$$

which implies that

$$u = \frac{v + v^*}{2} \qquad w = \frac{v - v^*}{2j}.$$

Making use of the matrix exponential, we have

$$e^{At}v == e^{\lambda t}(u + jw) = e^{\sigma t}\big((u\cos\omega t - w\sin\omega t) + j(u\sin\omega t + w\cos\omega t)\big),$$

which implies

$$\begin{aligned} e^{At}u &= \frac{1}{2}\Big(e^{At}v + e^{At}v^*\Big) = ue^{\sigma t}\cos\omega t - we^{\sigma t}\sin\omega t \\ e^{At}w &= \frac{1}{2j}\Big(e^{At}v - e^{At}v^*\Big) = ue^{\sigma t}\sin\omega t + we^{\sigma t}\cos\omega t. \end{aligned}$$

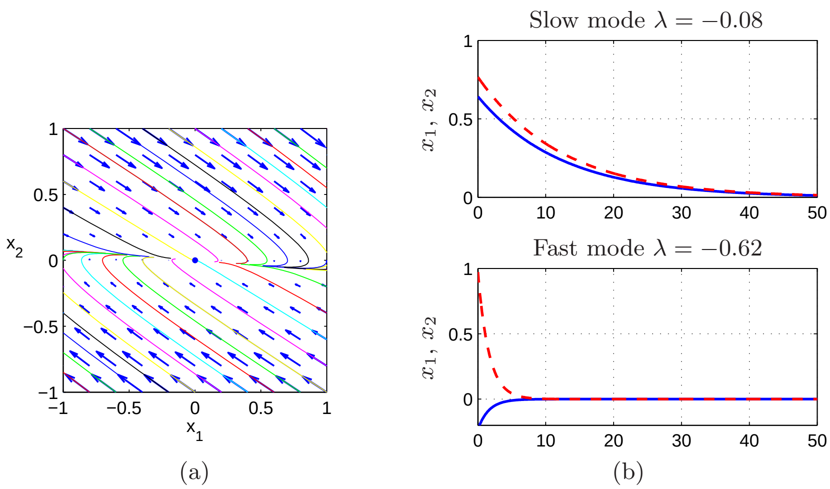


Figure 5.4: Illustration of the notion of modes for a second order system with real eigenvalues. The left figure (a) shows the phase plane and the modes corresponds to solutions that start on the eigenvectors. The time functions are shown in (b). The ratios of the states are also computed to show that they are constant for the modes.

A solution with initial conditions in the subspace spanned by the real part u and imaginary part v of the eigenvector will thus remain in that subspace. The solution will be logarithmic spiral characterized by σ and ω. We again call λ a mode of the system and v the mode shape.

If a matrix A has a n distinct eigenvalues $\lambda_1, \dots, \lambda_n$, then the initial condition response can be written as a linear combination of the modes. To see this, suppose for simplicity that we have all real eigenvalues with corresponding unit eigenvectors $v_1, \dots, v_n$. From linear algebra, these eigenvectors are linearly independent and we can write the initial condition $x(0)$ as

$$x(0) = \alpha_1 v_1 + \alpha_2 v_2 + \cdots \alpha_n v_n.$$

Using linearity, the initial condition response can be written as

$$x(t) = \alpha_1 e^{\lambda_1 t} v_1 + \alpha_2 e^{\lambda_2 t} v_2 + \cdots + \alpha_n e^{\lambda_n t} v_n.$$

Thus, the response is a linear combination the modes of the system, with the amplitude of the individual modes growing or decaying as $e^{\lambda_i t}$. The case for distinct complex eigenvalues follows similarly (the case for non-distinct eigenvalues is more subtle and is described in the section on the Jordan form, below).

Linear Input/Output Response

We now return to the general input/output case in equation (5.3), repeated here:

$$\begin{aligned}\frac{dx}{dt} &= Ax + Bu\\ y &= Cx + Du.\end{aligned} \tag{5.8}$$

Using the matrix exponential, the solution to equation (5.8) can be written as follows.

Theorem 5.3. *The solution to the linear differential equation* (5.8) *is given by*

$$x(t) = e^{At}x(0) + \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau. \tag{5.9}$$

Proof. To prove this, we differentiate both sides and use the property (5.7) of the matrix exponential. This gives

$$\frac{dx}{dt} = Ae^{At}x(0) + \int_0^t Ae^{A(t-\tau)}Bu(\tau)d\tau + Bu(t) = Ax + Bu,$$

which proves the result. Notice that the calculation is essentially the same as for proving the result for a first order equation. □

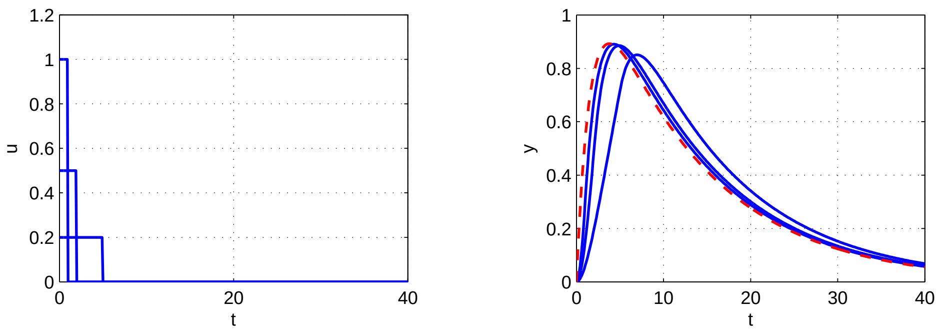
It follows from equations (5.8) and (5.9) that the input/output relation for a linear system is given by

$$y(t) = Ce^{At}x(0) + \int_0^t Ce^{A(t-\tau)}Bu(\tau)d\tau + Du(t). \tag{5.10}$$

It is easy to see from this equation that the output is jointly linear in both the initial conditions and the state: this follows from the linearity of matrix/vector multiplication and integration.

Equation (5.10) is called the *convolution equation* and it represents the general form of the solution of a system of coupled linear differential equations. We see immediately that the dynamics of the system, as characterized by the matrix A, play a critical role in both the stability and performance of the system. Indeed, the matrix exponential describes *both* what happens when we perturb the initial condition and how the system responds to inputs.

Another interpretation of the convolution equation can be given using the



(a) (b)

Figure 5.5: (a) Pulses of width 5, 2 and 1, each with total area equal to 1. (b) The pulse responses (solid) and impulse response (dashed) for a linear system with eigenvalues $\lambda = \{-0.08, -0.62\}$.

concept of the *impulse response* of a system. Consider the application of an input signal $u(t)$ given by the following equation:

$$u(t) = p_\epsilon(t) = \begin{cases} 0 & t < 0 \\ 1/\epsilon & 0 \le t < \epsilon \\ 0 & t \ge \epsilon. \end{cases} \tag{5.11}$$

This signal is a "pulse" of duration ϵ and amplitude $1/\epsilon$, as illustrated in Figure 5.5a. We define an *impulse*, $\delta(t)$, to be the limit of this signal as $\epsilon \to 0$:

$$\delta(t) = \lim_{\epsilon \to 0} p_\epsilon(t). \tag{5.12}$$

This signal, sometimes called a *delta function,* is not physically achievable but provides a convenient abstraction for understanding the response of a system. Note that the integral of an impulse is a unit step function, sometimes written as $1(t)$:

$$\begin{aligned} 1(t) &= \int_0^t \delta(\tau)\, d\tau = \int_0^t \lim_{\epsilon \to 0} p_\epsilon(t)\, d\tau \qquad && t > 0 \\ &= \lim_{\epsilon \to 0} \int_0^t p_\epsilon(t)\, d\tau = \lim_{\epsilon \to 0} \int_0^\epsilon 1/\epsilon\, d\tau = 1 && t > 0 \end{aligned}$$

In particular, the integral of an impulse over an arbitrarily short period of time is identically 1.

We define the *impulse response* of a system, $h(t)$, to be the output corresponding to an impulse as its input:

$$h(t) = \int_0^t Ce^{A(t-\tau)} B\delta(\tau)\, d\tau = Ce^{At}B, \tag{5.13}$$

where the second equality follows from the fact that $\delta(t)$ is zero everywhere except the origin and its integral is identically one. We can now write the convolution equation in terms of the initial condition response and the convolution of the impulse response and the input signal,

$$y(t) = Ce^{At}x(0) + \int_0^t h(t-\tau)u(\tau)\, d\tau. \tag{5.14}$$

One interpretation of this equation, explored in Exercise 6, is that the response of the linear system is the superposition of the response to an infinite set of shifted impulses whose magnitude is given by the input, $u(t)$. Note that the second term in this equation is identical to equation (5.4) and it can be shown that the impulse response is formally equivalent to the derivative of the step response.

The use of pulses as an approximation of the impulse response provides a mechanism for identifying the dynamics of a system from data. Figure 5.5b shows the pulse responses of a system for different pulse widths. Notice that the pulse responses approaches the impulse response as the pulse width goes to zero. As a general rule, if the fastest eigenvalue of a stable system has real part $-\lambda_{\max}$, then a pulse of length ϵ will provide a good estimate of the impulse response if $\epsilon\lambda_{\max} < 1$. Note that for Figure 5.5, a pulse width of $\epsilon = 1$ s gives $\epsilon\lambda_{\max} = 0.62$ and the pulse response is very close to the impulse response.

Coordinate Changes

The components of the input vector u and the output vector y are unique physical signals, but the state variables depend on the coordinate system chosen to represent the state. The choice of coordinates affects the values of the matrices A, B and C that are used in the model. (The direct term D is not affecting since it maps inputs to outputs.) We now investigate some of the consequences of changing coordinate systems.

Introduce new coordinates z by the transformation $z = Tx$, where T is an invertible matrix. It follows from equation (5.3) that

$$\begin{aligned}\frac{dz}{dt} &= T(Ax + Bu) = TAT^{-1}z + TBu = \tilde{A}z + \tilde{B}u\\ y &= Cx + DU \quad\;\; = CT^{-1}z + Du \quad\;\; = \tilde{C}z + Du.\end{aligned}$$

The transformed system has the same form as equation (5.3) but the matrices A, B and C are different:

$$\tilde{A} = TAT^{-1}, \qquad \tilde{B} = TB, \qquad \tilde{C} = CT^{-1}, \qquad \tilde{D} = D. \tag{5.15}$$

As we shall see in several places later in the text, there are often special choices of coordinate systems that allow us to see a particular property of the system, hence coordinate transformations can be used to gain new insight into the dynamics.

We can also compare the solution of the system in transformed coordinates to that in the original state coordinates. We make use of an important property of the exponential map,

$$e^{TST^{-1}} = Te^{S}T^{-1},$$

which can be verified by substitution in the definition of the exponential map. Using this property, it is easy to show that

$$x(t) = T^{-1}z(t) = T^{-1}e^{\tilde{A}t}Tx(0) + T^{-1}\int_0^t e^{\tilde{A}(t-\tau)}\tilde{B}u(\tau)\,d\tau.$$

From this form of the equation, we see that if it is possible to transform A into a form $\tilde{A}$ for which the matrix exponential is easy to compute, we can use that computation to solve the general convolution equation for the untransformed state x by simple matrix multiplications. This technique is illustrated in the next section.

Example 5.5 (Modal form)**.** Suppose that A has n real, distinct eigenvalues, $\lambda_1, \dots, \lambda_n$. It follows from matrix linear algebra that the corresponding eigenvectors $v_1, \dots v_n$ are linearly independent and form a basis for $\mathbb{R}^n$. Suppose that we transform coordinates according to the rule

$$x = Mz \qquad M = \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix}.$$

Setting $T = M^{-1}$, it is easy to show that

$$\tilde{A} = TAT^{-1} = M^{-1}AM = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}.$$

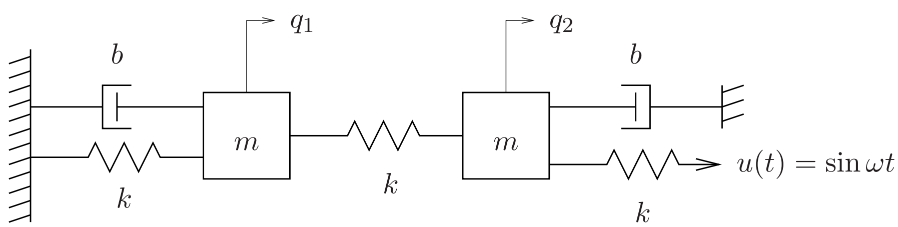


Figure 5.6: Coupled spring mass system.

To see this, note that if we multiple $M^{-1}AM$ by the basis elements

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \dots \qquad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

we get precisely $\lambda_i e_i$, which is the same as multiplying the diagonal form by the canonical basis elements. Since this is true for each e_i, $i = 1, \dots, n$ and since the these vectors form a basis for $\mathbb{R}^n$, the transformed matrix must be in the given form. This is precisely the diagonal form of Example 5.4, which is also called the *modal form* for the system. ∇

Example 5.6 (Coupled mass spring system)**.** Consider the coupled mass spring system shown in Figure 5.6. The input to this system is the sinusoidal motion of the end of rightmost spring and the output is the position of each mass, q_1 and q_2. The equations of motion for the system are given by

$$\begin{aligned} m_1\ddot{q}_1 &= -2kq_1 - c\dot{q}_1 + kq_2 \\ m_2\ddot{q}_2 &= kq_1 - 2kq_2 - c\dot{q}_2 + ku \end{aligned}$$

In state-space form, we define the state to be $x = (q_1, q_2, \dot{q}_1, \dot{q}_2)$ and we can rewrite the equations as

$$\dot{x} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{2k}{m} & \frac{k}{m} & -\frac{c}{m} & 0 \\ \frac{k}{m} & -\frac{2k}{m} & 0 & -\frac{c}{m} \end{pmatrix} x + \begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{k}{m} \end{pmatrix} u.$$

This is a coupled set of four differential equations and quite difficult to solve in analytical form.

We now define a transformation $z = Tx$ that puts this system into a simpler form. Let $z_1 = \frac{1}{2}(q_1 + q_2)$, $z_2 = \dot{z}_1$, $z_3 = \frac{1}{2}(q_1 - q_2)$ and $z_4 = \dot{z}_3$, so that

$$z = Tx = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} x.$$

Using the coordinate transformations described above (or simple substitution of variables, which is equivalent), we can write the system in the z coordinates as

$$\dot{z} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\frac{k}{m} & -\frac{c}{m} & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\frac{3k}{m} & -\frac{c}{m} \end{pmatrix} x + \begin{pmatrix} 0 \\ \frac{k}{2m} \\ 0 \\ -\frac{k}{2m} \end{pmatrix} u.$$

Note that the resulting matrix equations are are block diagonal and hence decoupled. We can thus solve for the solutions by computing the two sets of second order system represented by the states (z_1, z_2) and (z_3, z_4). Indeed, the functional form of each set of equations is identical to that of a single spring mass system (Section 2.1).

Once we have solved the two sets of independent second order equations, we can recover the dynamics in the original coordinates by inverting the state transformation and writing $x = T^{-1}z$. We can also determine the stability of the system by looking at the stability of the independent second order systems (Exercise 1). ∇

5.3 Stability and Performance

The special form of a linear system and its solution through the convolution equation allow us to analytically solve for the stability of equilibrium points and input/output performance properties.

Stability of Linear Systems

For a linear system, the stability of the equilibrium point at the origin can be determined by looking at the eigenvalues of the stability matrix A:

$$\lambda(A) = \{s \in \mathbb{C} : \det(sI - A) = 0\}.$$

We use the notation λ_i for the ith eigenvalue of A, so that $\lambda_i \in \lambda(A)$.

The easiest class of linear systems to analyze are those whose system matrices are in diagonal form. In this case, the dynamics have the form

$$\frac{dx}{dt} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} x + \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} u$$
$$y = \begin{pmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_n \end{pmatrix} x + Du.$$

Using Example 5.4, it is easy to show that the state trajectories for this system are independent of each other, so that we can write the solution in terms of n individual systems

$$\dot{x}_i = \lambda_i x_i + \beta_i u.$$

Each of these scalar solutions is of the form

$$x_i(t) = e^{\lambda_i t} x(0) + \int_0^t e^{\lambda(t-\tau)} u(t)\, dt.$$

If we consider the stability of the system when $u = 0$, we see that the equilibrium point $x_e = 0$ is stable if $\lambda_i \leq 0$ and asymptotically stable if $\lambda_i < 0$.

Very few systems are diagonal, but some systems can be transformed into diagonal form via coordinate transformations. One such class of systems is those for which the dynamics matrix has distinct (non-repeating) eigenvalues, as outlined in Example 5.5. In this case it is possible to find a matrix T such that the matrix TAT^{-1} and the transformed system is in diagonal form, with the diagonal elements equal to the the eigenvalues of the original matrix A. We can reason about the stability of the original system by noting that $x(t) = T^{-1}z(t)$ and so if the transformed system is stable (or asymptotically stable) then the original system has the same type stability.

For more complicated systems, we make use of the following theorem, proved in the next section:

Theorem 5.4. *The system*

$$\dot{x} = Ax$$

is asymptotically stable if and only if all eigenvalues of A all have strictly negative real part and is unstable if any eigenvalue of A has strictly positive real part.

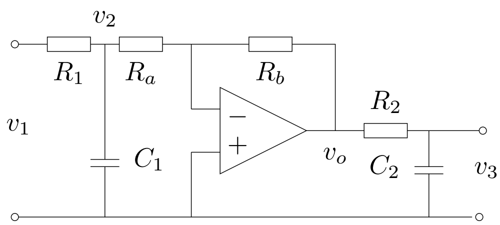


Figure 5.7: Active filter circuit using an operational amplifier.

Example 5.7 (Active filter)**.** Consider the op amp circuit shown in Figure 5.7. There are two energy storage elements, the capacitors C_1 and C_2. We choose their voltages, v_2 and v_3, as states. The dynamics for the system (Chapter 3, Exercise 5) are given by

$$\begin{aligned} \dot{x} &= \begin{pmatrix} -\frac{1}{R_1C_1} - \frac{1}{R_aC_1} & 0 \\ \frac{R_b}{R_a}\frac{1}{R_2C_2} & -\frac{1}{R_2C_2} \end{pmatrix} x + \begin{pmatrix} \frac{1}{R_1C_1} \\ 0 \end{pmatrix} u \\ y &= \begin{pmatrix} 0 & 1 \end{pmatrix} x, \end{aligned}$$

where $u = v_1$ and $y = v_3$. The eigenvalues of the dynamics matrix, A, are

$$\lambda_1 = -\frac{1}{R_1C_1} - \frac{1}{R_aC_1} \qquad \lambda_2 = -\frac{1}{R_2C_2}.$$

Assuming all capacitances and resistances are positive, these eigenvalues are both real and negative, and hence the equilibrium point at $x = 0$ is asymptotically stable. This implies, in particular, that if no input voltage is applied, the voltages around the system will all converge to zero as $t \to \infty$. ∇

Jordan Form

Some matrices with equal eigenvalues cannot be transformed to diagonal form. They can however be transformed to the Jordan form. In this form the dynamics matrix has the eigenvalues along the diagonal. When there are equal eigenvalues there may be ones appearing in the super diagonal indicating that there is coupling between the states.

More specifically, we define a matrix to be in *Jordan form* if it can be written as

$$J = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & & 0 \\ 0 & \dots & \ddots & 0 \\ 0 & \dots & & J_k \end{pmatrix} \quad \text{where} \quad J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_i & 1 \\ 0 & \dots & 0 & 0 & \lambda_i \end{pmatrix}. \tag{5.16}$$

Each matrix J_i is called a *Jordan block* and λ_i for that block corresponds to an eigenvalue of J.

Theorem 5.5 (Jordan decomposition)**.** *Any matrix $A \in \mathbb{R}^{n\times n}$ can be transformed into Jordan form with the eigenvalues of A determining λ_i in the Jordan form.*

Proof. See any standard text on linear algebra, such as Strang [Str88]. □

Converting a matrix into Jordan form can be very complicated, although MATLAB can do this conversion for numerical matrices using the `Jordan` function. The structure of the resulting Jordan form is particularly interesting since there is no requirement that the individual λ_i's be unique, and hence for a given eigenvalue we can have one or more Jordan blocks of different size. We say that a Jordan block J_i is *trivial* if J_i is a scalar (1×1 block).

Once a matrix is in Jordan form, the exponential of the matrix can be computed in terms of the Jordan blocks:

$$e^J = \begin{pmatrix} e^{J_1} & 0 & \dots & 0 \\ 0 & e^{J_2} & & 0 \\ 0 & \dots & \ddots & 0 \\ 0 & \dots & & e^{J_k}. \end{pmatrix} \tag{5.17}$$

This follows from the block diagonal form of J. The exponentials of the Jordan blocks can in turn be written as

$$e^{J_i t} = \begin{pmatrix} e^{\lambda_i t} & t\,e^{\lambda_i t} & \frac{t^2}{2!}\,e^{\lambda_i t} & \dots & \frac{t^{n-1}}{(n-1)!}\,e^{\lambda_i t} \\ 0 & e^{\lambda_i t} & t\,e^{\lambda_i t} & \dots & \frac{t^{n-2}}{(n-2)!}\,e^{\lambda_i t} \\ & & e^{\lambda_i t} & \ddots & \\ & & & \ddots & t\,e^{\lambda_i t} \\ 0 & & & & e^{\lambda_i t} \end{pmatrix} \tag{5.18}$$

When there are multiple eigenvalues, the invariant subspaces representing the modes correspond to the Jordan blocks of the matrix A . Note that λ may be complex, in which case the transformation T that converts a matrix into Jordan form will also be complex. When λ has a non-zero imaginary component, the solutions will have oscillatory components since

$$e^{\sigma+j\omega t} = e^{\sigma t}(\cos\omega t + j\sin\omega t).$$

We can now use these results to prove Theorem 5.4.

Proof of Theorem 5.4. Let $T \in \mathbb{C}^{n\times n}$ be an invertible matrix that transforms A into Jordan form, $J = TAT^{-1}$. Using coordinates $z = Tx$, we can write the solution $z(t)$ as

$$z(t) = e^{Jt}z(0).$$

Since any solution $x(t)$ can be written in terms of a solution $z(t)$ with $z(0) = Tx(0)$, it follows that it is sufficient to prove the theorem in the transformed coordinates.

The solution $z(t)$ can be written as a combination of the elements of the matrix exponential and from equation (5.18) these elements all decay to zero for arbitrary $z(0)$ if and only if $\operatorname{Re}\lambda_i < 0$. Furthermore, if any λ_i has positive real part, then there exists an initial condition $z(0)$ such that the corresponding solution increases without bound. Since we can scale this initial condition to be arbitrarily small, it follows that the equilibrium point is unstable if any eigenvalue has positive real part. □

The existence of a canonical form allows us to prove many properties of linear systems by changing to a set of coordinates in which the A matrix is in Jordan form. We illustrate this in the following proposition, which follows along the same lines as the proof of Theorem 5.4.

Proposition 5.6. *Suppose that the system*

$$\dot{x} = Ax$$

has no eigenvalues with strictly positive real part and one or more eigenvalues with zero real part. Then the system is stable if and only if the Jordan blocks corresponding to each eigenvalue with zero real part are scalar (1×1) *blocks.*

Proof. Exercise 3. □

Input/Output Response

So far, this chapter has focused on the stability characteristics of a system. While stability is often a desirably feature, stability alone may not be sufficient in many applications. We will want to create feedback systems that quickly react to changes and give high performance in measurable ways.

We return now to the case of an input/output state space system

$$\begin{aligned} \frac{dx}{dt} &= Ax + Bu \\ y &= Cx + Du, \end{aligned} \tag{5.19}$$

where $x \in \mathbb{R}^n$ is the state and $u, y \in \mathbb{R}$ are the input and output. The general form of the solution to equation (5.19) is given by the convolution equation:

$$y(t) = Ce^{At}x(0) + \int_0^t Ce^{A(t-\tau)}Bu(\tau)d\tau + Du(t).$$

We see from the form of this equation that the solution consists of an initial condition response and an input response.

The input response, corresponding to the second term in the equation above, itself consists of two components—the transient response and steady state response. The transient response occurs in the first period of time after the input is applied and reflects the mismatch between the initial condition and the steady state solution. The steady state response is the portion of the output response that reflects the long term behavior of the system under the given inputs. For inputs that are periodic, the steady state response will often also be periodic. An example of the transient and steady state response is shown in Figure 5.8.

Step Response

A particularly common form of input is a *step input*, which represents an abrupt change in input from one value to another. A *unit step* is defined as

$$u = 1(t) = \begin{cases} 0 & t = 0 \\ 1 & t > 0. \end{cases}$$

The *step response* of the system (5.3) is defined as the output $y(t)$ starting from zero initial condition (or the appropriate equilibrium point) and given a *step input*. We note that the step input is discontinuous and hence is not

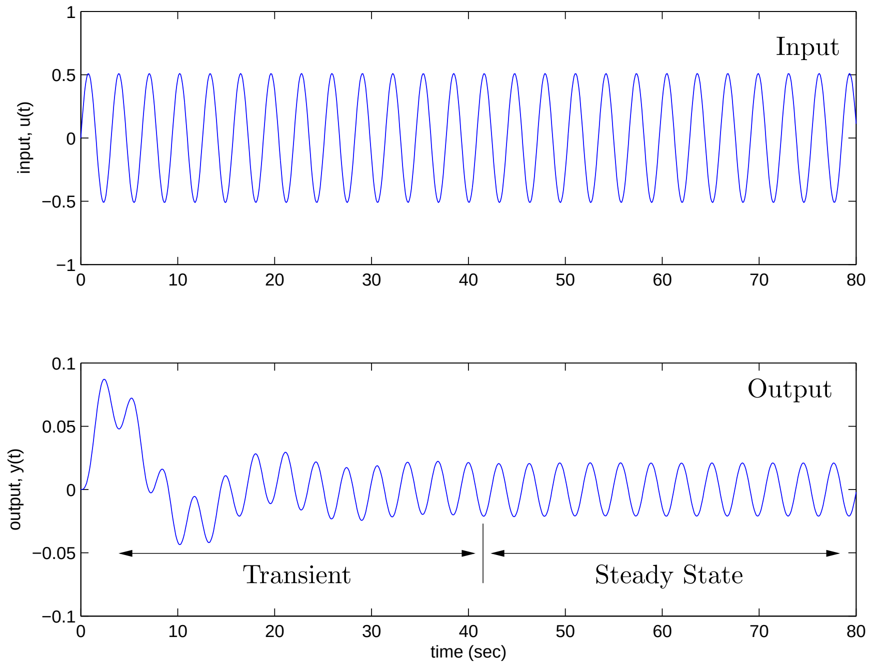


Figure 5.8: Transient versus steady state response. The top plot shows the input to a linear system and the bottom plot the corresponding output. The output signal initially undergoes a transient before settling into its steady state behavior.

physically implementable. However, it is a convenient abstraction that is widely used in studying input/output systems.

We can compute the step response to a linear system using the convolution equation. Setting $x(0) = 0$ and using the definition of the step input above, we have

$$\begin{aligned} y(t) &= \int_0^t Ce^{A(t-\tau)}Bu(\tau)d\tau + Du(t) \\ &= \int_0^t Ce^{A(t-\tau)}Bd\tau + D \qquad\qquad t > 0. \end{aligned}$$

If A has eigenvalues with negative real part (implying that the origin is a stable equilibrium point in the absence of any input), then we can rewrite the solution as

$$y(t) = \underbrace{CA^{-1}e^{At}B}_{\text{transient}} + \underbrace{D - CA^{-1}B}_{\text{steady state}} \qquad t > 0. \tag{5.20}$$

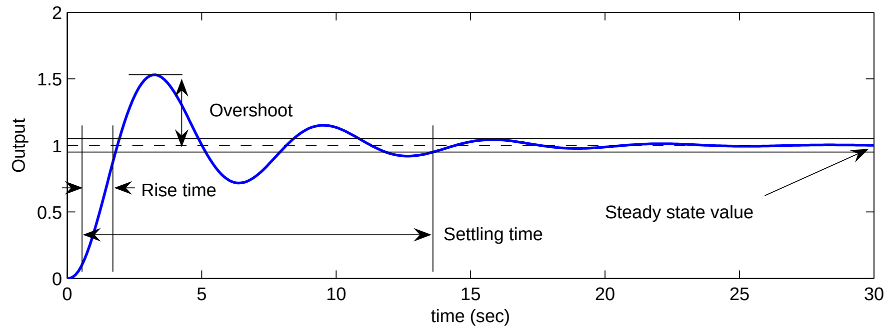


Figure 5.9: Sample step response

The first term is the transient response and decays to zero as $t \to \infty$. The second term is the steady state response and represents the value of the output for large time.

A sample step response is shown in Figure 5.9. Several terms are used when referring to a step response:

Steady state value The steady state value, y_{ss}, of a step response is the final level of the output, assuming it converges.

Rise time The rise time, T_r, is the amount of time required for the signal to go from 10% of its final value to 90% of its final value. It is possible to define other limits as well, but in this book we shall use these percentages unless otherwise indicated.

Overshoot The overshoot, M_p, is the percentage of the final value by which the signal initially rises above the final value. This usually assumes that future values of the signal do not overshoot the final value by more than this initial transient, otherwise the term can be ambiguous.

Settling time The settling time, T_s, is the amount of time required for the signal to stay within 5% of its final value for all future times. The settling time is also sometimes defined as reaching 1% or 2% of the final value (see Exercise 5).

In general these performance measures can depend on the amplitude of the input step, but for linear systems it can be shown that the quantities defined above are independent of the size of the step.

Frequency Response

The frequency response of an input/output system measures the way in which the system responds to a sinusoidal excitation on one of its inputs. As we have already seen for linear systems, the particular solution associated with a sinusoidal excitation is itself a sinusoid at the same frequency. Hence we can compare the magnitude and phase of the output sinusoid to the input. More generally, if a system has a sinusoidal output response at the same frequency as the input forcing, we can speak of the frequency response of the system.

To see this in more detail, we must evaluate the convolution equation (5.10) for $u = \cos \omega t$. This turns out to be a very messy computation, but we can make use of the fact that the system is linear to simplify the derivation. In particular, we note that

$$\cos \omega t = \frac{1}{2}\Big(e^{j\omega t} + e^{-j\omega t}\Big).$$

Since the system is linear, it suffices to compute the response of the system to the complex input $u(t) = e^{st}$ and we can always reconstruct the input to a sinusoid by averaging the responses corresponding to $s = j\omega t$ and $s = -j\omega t$.

Applying the convolution equation to the input $u = e^{st}$, we have

$$\begin{aligned} y(t) &= \int_0^t Ce^{A(t-\tau)}Be^{s\tau}d\tau + De^{st} \\ &= \int_0^t Ce^{A(t-\tau)+sI\tau}Bd\tau + De^{st} \\ &= e^{At}\int_0^t Ce^{(sI-A)\tau}Bd\tau + De^{st}. \end{aligned}$$

If we assume that none of the eigenvalues of A are equal to $s = \pm j\omega$, then the matrix $sI - A$ is invertible and we can write (after some algebra)

$$y(t) = \underbrace{Ce^{At}\Big(x(0) - (sI-A)^{-1}B\Big)}_{\text{transient}} + \underbrace{\Big(D + C(sI-A)^{-1}B\Big)e^{st}}_{\text{steady state}}.$$

Notice that once again the solution consists of both a transient component and a steady state component. The transient component decays to zero if the system is asymptotically stable and the steady state component is proportional to the (complex) input $u = e^{st}$.

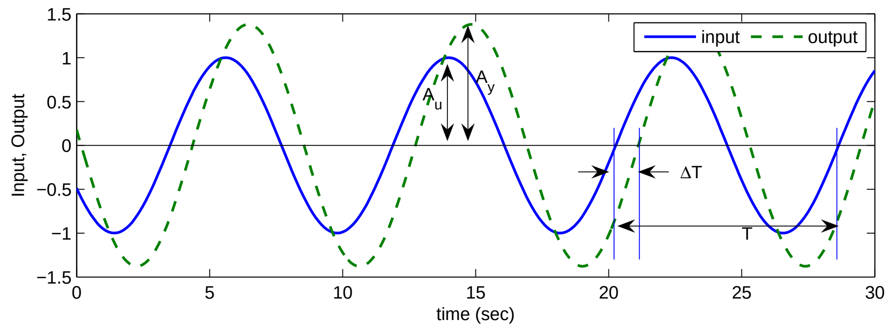


Figure 5.10: Frequency response, showing gain and phase. The phase lag is given by $\theta = -2\pi\Delta T/T$.

We can simplify the form of the solution slightly further by rewriting the steady state response as

$$y_{\text{ss}} = Me^{j\theta}e^{st} = Me^{(st+j\theta)}$$

where

$$Me^{j\theta} = C(sI - A)^{-1}B + D \tag{5.21}$$

and M and θ represent the magnitude and phase of the complex number $D + C(sI - A)^{-1}B$. When $s = j\omega$, we say that M is the *gain* and θ is the *phase* of the system at a given forcing frequency ω. Using linearity and combining the solutions for $s = +j\omega$ and $s = -j\omega$, we can show that if we have an input $u = A_u \sin(\omega t + \psi)$ and output $y = A_y \sin(\omega t + \varphi)$, then

$$\text{gain}(\omega) = \frac{A_y}{A_u} = M \qquad \text{phase}(\omega) = \varphi - \psi = \theta.$$

If the phase is positive, we say that the output "leads" the input, otherwise we say it "lags" the input.

A sample frequency response is illustrated in Figure 5.10. The solid line shows the input sinusoid, which has amplitude 1. The output sinusoid is shown as a dashed line, and has a different amplitude plus a shifted phase. The gain is the ratio of the amplitudes of the sinusoids, which can be determined by measuring the height of the peaks. The phase is determined by comparing the ratio of the time between zero crossings of the input and output to the overall period of the sinusoid:

$$\theta = -2\pi \cdot \frac{\delta T}{T}.$$

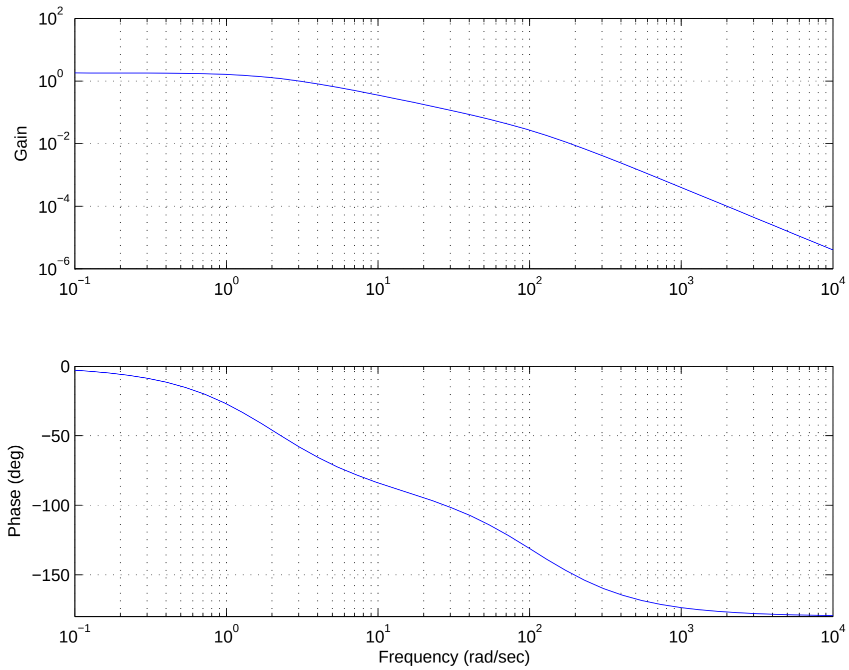


Figure 5.11: Frequency response for the active filter from Example 5.7. The upper plot shows the magnitude as a function of frequency (on a log-log scale) and the lower plot shows the phase (on a log-linear scale).

Example 5.8 (Active filter)**.** Consider the active filter presented in Example 5.7. The frequency response for the system can be computed using equation (5.21):

$$Me^{j\theta} = C(sI - A)^{-1}B + D = \frac{R_b/R_a}{(1 + R_2C_2s)(\frac{R_1+R_a}{R_a} + R_1C_1s)} \qquad s = j\omega.$$

The magnitude and phase are plotted in Figure 5.11 for $R_a = 1k\Omega$, $R_b = 100\ k\Omega$, $R_1 = 100\Omega$, $R_2 = 5k\Omega$ and $C_1 = C_2 = 100\ \mu$F. ∇

The gain at frequency $\omega = 0$ is called the *zero frequency gain* of the system and corresponds to the ratio between a constant input and the steady output:

$$M_0 = CA^{-1}B + D.$$

Note that the zero frequency gain is only well defined if A is invertible (and, in particular, if it does has not eigenvalues at 0). It is also important to note that the zero frequency gain is only a relevant quantity when a system is stable about the corresponding equilibrium point. So, if we apply a constant input $u = r$ then the corresponding equilibrium point

$$x_e = -A^{-1}Br$$

must be stable in order to talk about the zero frequency gain. (In electrical engineering, the zero frequency gain is often called the "DC gain". DC stands for "direct current" and reflects the common separation of signals in electrical engineering into a direct current (zero frequency) term and an alternating current (AC) term.)

5.4 Second Order Systems

One class of systems that occurs frequently in the analysis and design of feedback systems is second order, linear differential equations. Because of their ubiquitous nature, it is useful to apply the concepts of this chapter to that specific class of systems and build more intuition about the relationship between stability and performance.

The canonical second order system is a differential equation of the form

$$\begin{aligned} &\ddot{q} + 2\zeta\omega_0\dot{q} + \omega_0^2 q = ku \\ &y = q. \end{aligned} \tag{5.22}$$

In state space form, this system can be represented as

$$\begin{aligned} \dot{x} &= \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & -2\zeta\omega_0 \end{pmatrix} x + \begin{pmatrix} 0 \\ k \end{pmatrix} u \\ y &= \begin{pmatrix} 1 & 0 \end{pmatrix} x \end{aligned} \tag{5.23}$$

The eigenvalues of this system are given by

$$\lambda = -\zeta\omega_0 \pm \sqrt{\omega_0^2(\zeta^2 - 1)}$$

and we see that the origin is a stable equilibrium point if $\omega_0 > 0$ and $\zeta > 0$. Note that the eigenvalues are complex if $\zeta < 1$ and real otherwise. Equations (5.22) and (5.23) can be used to describe many second order systems, including a damped spring mass system and an active filter, as shown in the examples below.

The form of the solution depends on the value of ζ, which is referred to as the *damping factor* for the system. If $\zeta > 1$, we say that the system is *overdamped* and the natural response ($u = 0$) of the system is given by

$$y(t) = \frac{\beta x_{10} + x_{20}}{\beta - \alpha} e^{-\alpha t} - \frac{\alpha x_{10} + x_{20}}{\beta - \alpha} e^{-\beta t}$$

where $\alpha = \omega_0(\zeta + \sqrt{\zeta^2 - 1})$ and $\beta = \omega_0(\zeta - \sqrt{\zeta^2 - 1})$. We see that the response consists of the sum of two exponentially decaying signals. If $\zeta = 1$ then the system is *critically damped* and solution becomes

$$y(t) = e^{-\zeta\omega_0 t}\big(x_{10} + (x_{20} + \zeta\omega_0 x_{10})t\big).$$

Note that this is still asymptotically stable as long as $\omega_0 > 0$, although the second term in the solution is increasing with time (but more slowly than the decaying exponential that multiplies it).

Finally, if $0 < \zeta < 1$, then the solution is oscillatory and equation (5.22) is said to be *underdamped*. The parameter ω_0 is referred to as the natural frequency of the system, stemming from the fact that for small ζ, the eigenvalues of the system are approximately $\lambda = -\zeta \pm j\omega_0$. The natural response of the system is given by

$$y(t) = e^{-\zeta\omega_0 t}\left(x_{10}\cos\omega_d t + \Big(\frac{\zeta\omega_0}{\omega_d} x_{10} + \frac{1}{\omega_d} x_{20}\Big)\sin\omega_d t\right),$$

where $\omega_d = \omega_0\sqrt{1 - \zeta^2}$. For $\zeta \ll 1$, $\omega_d \approx \omega_0$ defines the oscillation frequency of the solution and ζ gives the damping rate relative to ω_0.

Because of the simple form of a second order system, it is possible to solve for the step and frequency responses in analytical form. The solution for the step response depends on the magnitude of ζ:

$$\begin{aligned}
y(t) &= \frac{k}{\omega_0^2}\left(1 - e^{-\zeta\omega_0 t}\cos\omega_d t + \frac{\zeta}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_0 t}\sin\omega_d t\right) && \zeta < 1 \\
y(t) &= \frac{k}{\omega_0^2}\big(1 - e^{-\omega_0 t}(1 + \omega_0 t)\big) && \zeta = 1 \\
y(t) &= \frac{k}{\omega_0^2}\left(1 - e^{-\omega_0 t} - \frac{1}{2(1+\zeta)} e^{\omega_0(1-2\zeta)t}\right) && \zeta > 1,
\end{aligned} \tag{5.24}$$

where we have taken $x(0) = 0$. Note that for the lightly damped case ($\zeta < 1$) we have an oscillatory solution at frequency ω_d, sometimes called the damped frequency.

The step responses of systems with $k = \omega^2$ and different values of ζ are shown in Figure 5.12, using a scaled time axis to allow an easier comparison.

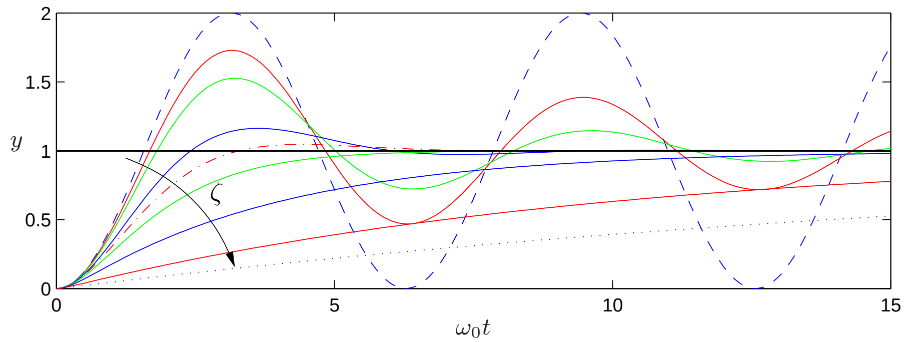


Figure 5.12: Normalized step responses h for the system (5.23) for $\zeta = 0$ (dashed), 0.1, 0.2, 0.5, 0.707 (dash dotted), 1, 2, 5 and 10 (dotted).

The shape of the response is determined by ζ and the speed of the response is determined by ω_0 (including in the time axis scaling): the response is faster if ω_0 is larger. The step responses have an overshoot of

$$M_p = \begin{cases} e^{-\pi\zeta/\sqrt{1-\zeta^2}} & \text{for } |\zeta| < 1 \\ 0 & \text{for } \zeta \geq 1. \end{cases} \tag{5.25}$$

For $\zeta < 1$ the maximum overshoot occurs at

$$t_{max} = \frac{\pi}{\omega_0\sqrt{1-\zeta^2}}. \tag{5.26}$$

The maximum decreases and is shifted to the right when ζ increases and it becomes infinite for $\zeta = 1$, when the overshoot disappears.

The frequency response can also be computed explicitly and is given by

$$Me^{j\theta} = \frac{\omega_0^2}{(j\omega)^2 + 2\zeta\omega_0(j\omega) + \omega_0^2} = \frac{\omega_0^2}{\omega_0^2 - \omega^2 + 2j\zeta\omega_0\omega}.$$

A graphical illustration of the frequency response is given in Figure 5.13. Notice the resonance peak that increases with decreasing ζ. The peak is often characterized by is Q-*value*, defined as $Q = 1/2\zeta$.

Example 5.9 (Damped spring mass)**.** The dynamics for a damped spring mass system are given by

$$m\ddot{q} + c\dot{q} + kq = u,$$

where m is the mass, q is the displacement of the mass, c is the coefficient of viscous friction, k is the spring constant and u is the applied force. We

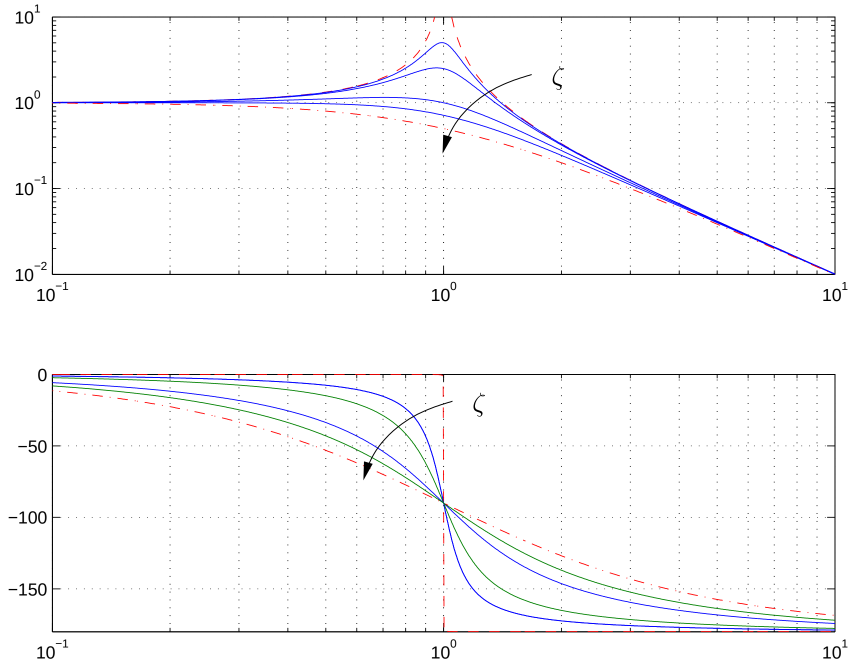


Figure 5.13: Frequency response of a the second order system (5.23). The upper curve shows the gain ratio, M, and the lower curve shows the phase shift, θ. The parameters is Bode plot of the system with $\zeta = 0$ (dashed), 0.1, 0.2, 0.5, 0.7 and 1.0 (dashed-dot).

can convert this into the standard second order for by dividing through by m, giving

$$\ddot{q} + \frac{c}{m}\dot{q} + \frac{k}{m}q = \frac{1}{m}u.$$

Thus we see that the spring mass system has natural frequency and damping ratio given by

$$\omega_0 = \sqrt{\frac{k}{m}} \qquad \zeta = \frac{c}{2\sqrt{km}}$$

(note that we have use the symbol k for the stiffness here; it should not be confused with the gain term in equation (5.22)). ∇

One of the other reasons why second order systems play such an important role in feedback systems is that even for more complicated systems the response is often dominated by the "dominant eigenvalues". To define these more precisely, consider a system with eigenvalues λ_i, $i = 1, \ldots, n$. We define the damping factor for a complex eigenvalue λ to be

$$\zeta = \frac{-\operatorname{Re}\lambda}{|\lambda|}$$

We say that a complex conjugate pair of eigenvalues λ, λ^* is a dominant pair if it has the lowest damping factor compared with all other eigenvalues of the system.

Assuming that a system is stable, the dominant pair of eigenvalues tends to be the most important element of the response. To see this, assume that we have a system in Jordan form with a simple Jordan block corresponding to the dominant pair of eigenvalues:

$$\begin{aligned}
\dot{z} &= \begin{pmatrix} \lambda & & & & \\ & \lambda^* & & & \\ & & J_2 & & \\ & & & \ddots & \\ & & & & J_k \end{pmatrix} z + Bu \\
y &= Cz.
\end{aligned}$$

(Note that the state z may be complex due to the Jordan transformation.) The response of the system will be a linear combination of the responses from each of the individual Jordan subsystems. As we see from Figure 5.12, for $\zeta < 1$ the subsystem with the slowest response is precisely the one with the smallest damping factor. Hence when we add the responses from each of the individual subsystems, it is the dominant pair of eigenvalues that will be dominant factor after the initial transients due to the other terms in the solution. While this simple analysis does not always hold (for example, if some non-dominant terms have large coefficients due to the particular form of the system), it is often the case that the dominant eigenvalues dominate the (step) response of the system. The following example illustrates the concept.

5.5 Linearization

As described in the beginning of the chapter, a common source of linear system models is through the *approximation* of a nonlinear system by a linear one. These approximations are aimed at studying the local behavior of a system, where the nonlinear effects are expected to be small. In this section we discuss how to locally approximate a system by its linearization and what can be said about the approximation in terms of stability. We begin with an illustration of the basic concept using the speed control example from Chapter 2.

Example 5.10 (Cruise control)**.** The dynamics for the cruise control system are derived in Section 3.1 and have the form

$$m\frac{dv}{dt} = \alpha_n u T(\alpha_n v) - mgC_r - \tfrac{1}{2}\rho C_v A v^2 - mg\sin\theta, \tag{5.27}$$

where the first term on the right hand side of the equation is the force generated by the engine and the remaining three terms are the rolling friction, aerodynamic drag and gravitational disturbance force. There is an equilibrium (v_e, u_e) when the force applied by the engine balances the disturbance forces.

To explore the behavior of the system near the equilibrium we will linearize the system. A Taylor series expansion of equation (5.27) around the equilibrium gives

$$\frac{d(v-v_e)}{dt} = a(v-v_e) - b_g(\theta-\theta_e) + b(u-u_e) \tag{5.28}$$

where

$$a = \frac{u_e\alpha_n^2 T'(\alpha_n v_e) - \rho C_v A v_e}{m} \qquad b_g = g\cos\theta_e \qquad b = \frac{\alpha_n T(\alpha_n v_e)}{m} \tag{5.29}$$

and terms of second and higher order have been neglected. For a car in fourth gear with $v_e = 25$ m/s, $\theta_e = 0$ and the numerical values for the car from Section 3.1, the equilibrium value for the throttle is $u_e = 0.1687$ and the model becomes

$$\frac{d(v-v_e)}{dt} = -0.0101(v-v_e) + 1.3203(u-u_e) - 9.8(\theta-\theta_e) \tag{5.30}$$

This linear model describes how small perturbations in the velocity about the nominal speed evolve in time.

Figure 5.14, which shows a simulation of a cruise controller with linear and nonlinear models, indicates that the differences between the linear and nonlinear models is not visible in the graph. ∇

Linear Approximation

To proceed more formally, consider a single input, single output nonlinear system

$$\begin{aligned} \frac{dx}{dt} &= f(x,u) \qquad && x\in\mathbb{R}^n, u\in\mathbb{R} \\ y &= h(x,u) && y\in\mathbb{R} \end{aligned} \tag{5.31}$$

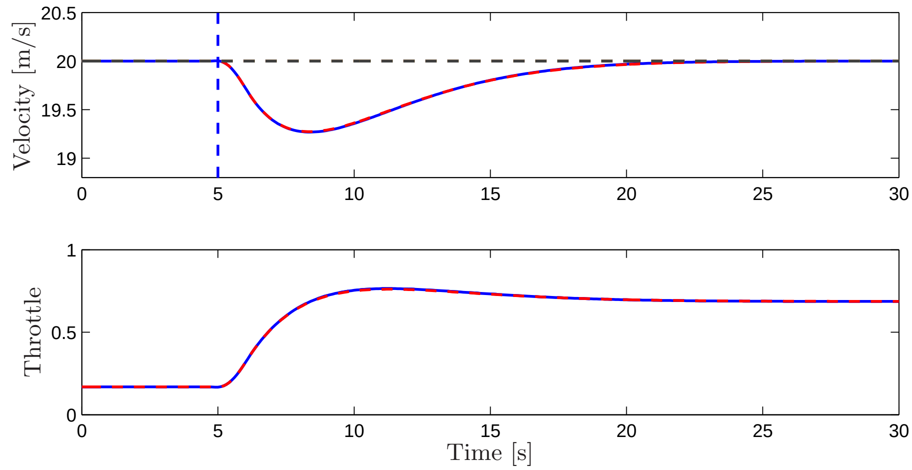


Figure 5.14: Simulated response of a vehicle with PI cruise control as it climbs a hill with a slope of 4°. The full lines is the simulation based on a nonlinear model and the dashed line shows the corresponding simulation using a linear model. The controller gains are $k_p = 0.5$ and $k_i = 0.1$.

with an equilibrium point at $x = x_e$, $u = u_e$. Without loss of generality, we assume that $x_e = 0$ and $u_e = 0$, although initially we will consider the general case to make the shift of coordinates explicit.

In order to study the *local* behavior of the system around the equilibrium point (x_e, u_e), we suppose that $x - x_e$ and $u - u_e$ are both small, so that nonlinear perturbations around this equilibrium point can be ignored compared with the (lower order) linear terms. This is roughly the same type of argument that is used when we do small angle approximations, replacing $\sin\theta$ with θ and $\cos\theta$ with 1 for θ near zero.

In order to formalize this idea, we define a new set of state variables z, inputs v, and outputs w:

$$z = x - x_e \qquad v = u - u_e \qquad w = y - h(x_e, u_e).$$

These variables are all close to zero when we are near the equilibrium point, and so in these variables the nonlinear terms can be thought of as the higher order terms in a Taylor series expansion of the relevant vector fields (assuming for now that these exist).

Example 5.11. Consider a simple scalar system,

$$\dot{x} = 1 - x^3 + u.$$

The point $(x_e, u_e) = (1, 0)$ is an equilibrium point for this system and we can thus set

$$z = x - 1 \qquad v = u.$$

We can now compute the equations in these new coordinates as

$$\begin{aligned}\dot{z} &= \frac{d}{dt}(x-1) = \dot{x} \\ &= 1 - x^3 + u = 1 - (z+1)^3 + v \\ &= 1 - z^3 - 3z^2 - 3z - 1 + v = -3z - 3z^2 - z^3 + v.\end{aligned}$$

If we now assume that x stays very close to the equilibrium point, then $z = x - x_e$ is small and $z \ll z^2 \ll z^3$. We can thus *approximate* our system by a *new* system

$$\dot{z} = -3z + v.$$

This set of equations should give behavior that is close to that of the original system as long as z remains small. ∇

More formally, we define the *Jacobian linearization* of the nonlinear system (5.31) as

$$\begin{aligned}\dot{z} &= Az + Bv \\ w &= Cz + Dv,\end{aligned} \tag{5.32}$$

where

$$\begin{aligned} A &= \left.\frac{\partial f(x,u)}{\partial x}\right|_{(x_e,u_e)} & B &= \left.\frac{\partial f(x,u)}{\partial u}\right|_{(x_e,u_e)} \\ C &= \left.\frac{\partial h(x,u)}{\partial x}\right|_{(x_e,u_e)} & D &= \left.\frac{\partial h(x,u)}{\partial u}\right|_{(x_e,u_e)} \end{aligned} \tag{5.33}$$

The system (5.32) approximates the original system (5.31) when we are near the equilibrium point that the system was linearized about.

It is important to note that we can only define the linearization of a system about an equilibrium point. To see this, consider a polynomial system

$$\dot{x} = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + u,$$

where $a_1 \neq 0$. There are a family of equilibrium points for this system given by $(x_e, u_e) = (x_e, -a_0 - a_1 x_e - a_2 x_e^2 - a_3 x_e^3)$ and we can linearize around any of these. Suppose that we try to linearize around the origin of the system, $x = 0$, $u = 0$. If we drop the higher order terms in x, then we get

$$\dot{x} = a_0 + a_1 x + u,$$

which is *not* the Jacobian linearization if $a_0 \neq 0$. The constant term must be kept and this is not present in (5.32). Furthermore, even if we kept the constant term in the approximate model, the system would quickly move away from this point (since it is "driven" by the constant term a_0) and hence the approximation could soon fail to hold.

Software for modeling and simulation frequently has facilities for performing linearization symbolically or numerically. The MATLAB command `trim` finds the equilibrium and `linmod` extracts linear state-space models from a SIMULINK system around an operating point.

Example 5.12 (Vehicle steering)**.** Consider the vehicle steering system introduced in Section 2.8. The nonlinear equations of motion for the system are given by equations (2.21)–(2.23) and can be written as

$$\frac{d}{dt}\begin{pmatrix} x \\ y \\ \theta \end{pmatrix} = \begin{pmatrix} v_0 \frac{\cos(\alpha+\theta)}{\cos\alpha} \\ v_0 \frac{\sin(\alpha+\theta)}{\cos\alpha} \\ \frac{v_0}{b}\tan\delta \end{pmatrix},$$

where x, y and θ are the position and orientation of the center of mass of the vehicle, v_0 is the velocity of the rear wheel, δ is the angle of the front wheel and α is the anglular devitation of the center of mass from the rear wheel along the instantaneous circle of curvature determined by the front wheel:

$$\alpha(\delta) = \arctan\Bigl(\frac{a\tan\delta}{b}\Bigr).$$

We are interested in the motion of the vehicle about a straight line path ($\theta = \theta_0$) with fixed velocity $v_0 \neq 0$. To find the relevant equilibrium point, we first set $\dot\theta = 0$ and we see that we must have $\delta = 0$, corresponding to the steering wheel being straight. This also yields $\alpha = 0$. Looking at the first two equations in the dynamics, we see that the motion in the xy direction is by definition *not* at equilibrium since $\dot x^2 + \dot y^2 = v_0^2 \neq 0$. Therefore we cannot formally linearize the full model.

Suppose instead that we are concerned with the lateral deviation of the vehicle from a straight line. For simplicity, we let $\theta_0 = 0$, which corresponds to driving along the x axis. We can then focus on the equations of motion in the y and θ directions, for which we have

$$\frac{d}{dt}\begin{pmatrix} y \\ \theta \end{pmatrix} = \begin{pmatrix} v_0 \frac{\sin(\alpha+\theta)}{\cos\alpha} \\ \frac{v_0}{b}\tan\delta \end{pmatrix}.$$

Abusing notation, we write $x = (y, \theta)$ and $u = \delta$ so that

$$f(x,u) = \begin{pmatrix} v_0 \frac{\sin(\alpha(u)+x_2)}{\cos\alpha(u)} \\ \frac{v_0}{b}\tan u, \end{pmatrix}$$

where the equilibrium point of interest is now given by $x = (0, 0)$ and $u = 0$.

To compute the linearization the model around the equilibrium point, we make use of the formulas (5.33). A straightforward calculation yields

$$A = \left.\frac{\partial f(x,u)}{\partial x}\right|_{\substack{x=0\\u=0}} = \begin{pmatrix} 0 & v_0 \\ 0 & 0 \end{pmatrix} \delta \qquad B = \left.\frac{\partial f(x,u)}{\partial u}\right|_{\substack{x=0\\u=0}} = \begin{pmatrix} v_0 \frac{a}{b} \\ \frac{v_0}{b} \end{pmatrix}$$

and the linearized system

$$\dot{z} = Az + Bv \tag{5.34}$$

thus provides an approximation to the original nonlinear dynamics.

A model can often be simplified further by introducing normalized dimension free variables. For this system, we can normalize lengths by the wheel base b and introduce a new time variable $\tau = v_0 t/b$. The time unit is thus the time is takes for the vehicle to travel one wheel base. We similarly normalize the lateral position and write $w_1 = y/b$, $w_2 = \theta$. The model (5.34) then becomes

$$\begin{aligned} \frac{dw}{d\tau} &= \begin{pmatrix} w_2 + \alpha v \\ v \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} w + \begin{pmatrix} \alpha \\ 1 \end{pmatrix} v \\ y &= \begin{pmatrix} 1 & 0 \end{pmatrix} w \end{aligned} \tag{5.35}$$

The normalized linear model for vehicle steering with non-slipping wheels is thus a linear system with only one parameter $\alpha = a/b$. ∇

Feedback Linearization

Another type of linearization is the use of feedback to convert the dynamics of a nonlinear system into a linear one. We illustrate the basic idea with an example.

Example 5.13 (Cruise control)**.** Consider again the cruise control system from Example 5.10, whose dynamics is given in equation (5.27). If we choose u as a feedback law of the form

$$u = \frac{1}{\alpha_n T(\alpha_n v)} \left(u' + mgC_r + \frac{1}{2}\rho C_v A v^2 \right) \tag{5.36}$$

then the resulting dynamics become

$$m\frac{dv}{dt} = u' + d \tag{5.37}$$

where $d = mg\sin\theta$ is the disturbance force due the slope of the road. If we now define a feedback law for u' (such as a PID controller), we can use equation (5.36) to compute the final input that should be commanded.

Equation (5.37) is a linear differential equation. We have essentially "inverted out" the nonlinearity through the use of the feedback law (5.36). This requires that we have an accurate measurement of the vehicle velocity v as well as an accurate model of the torque characteristics of the engine, gear ratios, drag and friction characteristics and mass of the car. While such a model is not generally available (remembering that the parameter values can change), if we design a good feedback law for u', then we can achieve robustness to these uncertainties. ∇

More generally, we say that a system of the form

$$\begin{aligned}\frac{dx}{dt} &= f(x,u)\\ y &= h(x)\end{aligned}$$

is *feedback linearizable* if we can find a control law $u = \alpha(x, v)$ such that the resulting closed loop system is input/output linear with input v and output u. To fully characterize such systems is beyond the scope of this text, but we note that in addition to changes in the input, we must also allow for (nonlinear) changes in the states that are used to describe the system, keeping only the input and output variables fixed. More details of this process can be found in the the textbooks by Isidori [Isi89] and Khalil [Kha92].

One case the comes up relatively frequently, and is hence worth special mention, is the set of mechanical systems of the form

$$M(q)\ddot{q} + C(q,\dot{q})\dot{q} + N(q,\dot{q}) = B(q)u.$$

Here $q \in \mathbb{R}^n$ is the configuration of the mechanical system, $M(q) \in \mathbb{R}^{n\times n}$ is the configuration-dependent inertia matrix, $C(q,\dot{q})\dot{q} \in \mathbb{R}^n$ represents the Coriolis forces, $N(q,\dot{q}) \in \mathbb{R}^n$ are additional nonlinear forces (such as stiffness and friction) and $B(q) \in \mathbb{R}^{n\times p}$ is the input matrix. If $p = n$ then we have the same number of inputs and configuration variables and if we further have that $B(q)$ is an invertible matrix for all configurations q, then we can choose

$$u = B^{-1}(q)\left(M(q)v - C(q,\dot{q})\dot{q} - N(q,\dot{q})\right). \tag{5.38}$$

The resulting dynamics become

$$M(q)\ddot{q} = M(q)v \qquad \Longrightarrow \qquad \ddot{q} = v,$$

which is a linear system. We can now use the tools of linear systems theory to analyze and design control laws for the linearized system, remembering to apply equation (5.38) to obtain the actual input that will be applied to the system.

This type of control is common in robotics, where it goes by the name of *computed torque*, and aircraft flight control, where it is called *dynamic inversion*.

Local Stability of Nonlinear Systems

Having constructed a linearized model around an equilibrium point, we can now ask to what extent this model predicts the behavior of the original nonlinear system. The following theorem gives a partial answer for the case of stability.

Theorem 5.7. *Consider the system* (5.31) *and let* $A \in \mathbb{R}^{n\times n}$ *be defined as in equations* (5.32) *and* (5.33)*. If the real part of the eigenvalues of* A *are strictly less than zero, then* x_e *is a locally asymptotically stable equilibrium point of* (5.31)*.*

This theorem shows that *global* asymptotic stability of the linearization implies *local* asymptotic stability of the original nonlinear system. The estimates provided by the proof of the theorem can be used to give a (conservative) bound on the domain of attraction of the origin. Systematic techniques for estimating the bounds on the regions of attraction of equilibrium points of nonlinear systems is an important area of research and involves searching for the "best" Lyapunov functions.

The proof of this theorem is the beyond the scope of this text, but can be found in [Kha92].

5.6 Further Reading

The idea to characterize dynamics by considering the responses to step inputs is due to Heaviside. The unit step is therefore also called the *Heaviside* *step function*. The majority of the material in this chapter is very classical and can be found in most books on dynamics and control theory, including early works on control such as James, Nichols and Phillips [JNP47], and more recent textbooks such as Franklin, Powell and Emami-Naeni [FPEN05] and Ogata [Oga01]. The material on feedback linearization is typically presented in books on nonlinear control theory, such as Khalil [Kha92]. Tracer methods are described in [She62].

5.7 Exercises

1. Compute the full solution to the couple spring mass system in Example 5.6 by transforming the solution for the block diagonal system back into the original set of coordinates. Show that the system is asymptotically stable if m, b and k are all greater than zero.

2. Using the computation for the matrix exponential, show that equation (5.18) holds for the case of a 3×3 Jordan block. (Hint: decompose the matrix into the form $S + N$ where S is a diagonal matrix.)

3. Prove Proposition 5.6.

4. Show that the step response for an asymptotically stable linear system is given by equation (5.20).

5. Consider a first order system of the form

$$\begin{aligned} \dot{x} &= -\tau x + u \\ y &= x. \end{aligned}$$

   We say that the parameter τ is the *time constant* for the system since the zero input system approaches the origin as $e^{\tau t}$. For a first order system of this form, show that the rise time of the system is approximately 2τ, a 5% settling time corresponds to approximately 3τ and a 2% settling time corresponds to approximately 4τ.

6. Show that a signal $u(t)$ can be decomposed in terms of the impulse function $\delta(t)$ as

$$u(t) = \int_0^t \delta(t - \tau) u(\tau)\, d\tau$$

   and use this decomposition plus the principle of superposition to show that the response of a linear system to an input $u(t)$ (assuming zero initial condition) can be written as

$$y(t) = \int_0^t h(t-\tau)u(\tau)\,d\tau,$$

where $h(t)$ is the impulse response of the system.

7. Consider a linear discrete time system of the form

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k \\ y_k &= Cx_k + Du_k. \end{aligned}$$

(a) Show that the general form of the output of a discrete time linear system is given by the discrete time convolution equation:

$$y_k = CA^k x_0 + \sum_{i=0}^{k} CA^i Bu_i + Du_k$$

(b) Show that a discrete time linear system is asymptotically stable if and only if all eigenvalues of A have magnitude strictly less than 1.

(c) Let $u_k = A\sin(\omega k)$ represent an oscillatory input with frequency $\omega < \pi$ (to avoid "aliasing"). Show that the steady state component of the response has gain M and phase θ where

$$Me^{j\theta} = C(j\omega I - A)^{-1}B + D.$$

(d) Show that if we have a nonlinear discrete time system

$$\begin{aligned} x_k &= f(x_k, u_k) \qquad && x_k \in \mathbb{R}^n, u \in \mathbb{R} \\ y_k &= h(x_k, u)k) \qquad && y \in \mathbb{R} \end{aligned}$$

then we can linearize the system around an equilibrium point (x_e, u_e) by defining the matrices A, B, C and D as in equation (5.33).

8. Consider the consensus protocol introduced in Example 2.13. Show that if the connectivity graph of the sensor network is connected, then we can find a gain γ such that the agent states converge to the average value of the measure quantity.